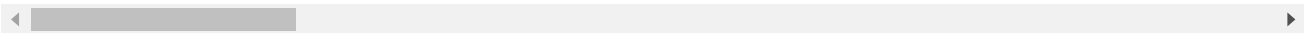# Feature Selection

Choosing the most effective variables

	gameID	leagueID	season	date	homeTeamID	awayTeamID	home_Goals	away_Goals
5437	6185	5	2014	2014-08-08 19:30:00	177	161	2	2
5438	6186	5	2014	2014-08-09 20:00:00	162	164	3	3
5439	6187	5	2014	2014-08-09 20:00:00	209	165	0	3
5440	6188	5	2014	2014-08-09 20:00:00	169	175	0	2
5441	6189	5	2014	2014-08-09 20:00:00	160	180	0	0

5 rows × 107 columns

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 12680 entries, 5437 to 12679
Data columns (total 107 columns):
 #     Column                                    Dtype
---    ------                                    -----
 0     gameID                                    int64
 1     leagueID                                  int64
 2     season                                    int64
 3     date                                      datetime64[ns]
 4     homeTeamID                                int64
 5     awayTeamID                                int64
 6     home_Goals                                int64
 7     away_Goals                                int64
 8     home_GoalsHalfTime                        int64
 9     away_GoalsHalfTime                        int64
 10    home_xGoals                               float64
 11    home_shots                                float64
 12    home_shotsOnTarget                        int64
 13    home_deep                                 float64
 14    home_ppda                                 float64
 15    home_fouls                                float64
 16    home_corners                              float64
 17    home_yellowCards                          int64
 18    home_redCards                             int64
 19    home_total_assists                        int64
 20    home_total_xAssists                       float64
 21    home_total_key_passes                     float64
 22    home_total_xGoalsChain                    float64
 23    home_total_xGoalsBuildup                  float64
 24    home_total_yellow_cards                   int64
 25    home_total_red_cards                      int64
 26    home_total_blocked_shots                  float64
 27    home_total_saved_shots                    float64
 28    away_xGoals                               float64
 29    away_shots                                float64
 30    away_shotsOnTarget                        int64
 31    away_deep                                 float64
 32    away_ppda                                 float64
 33    away_fouls                                float64
 34    away_corners                              int64
 35    away_yellowCards                          int64
 36    away_redCards                             int64
 37    away_total_assists                        int64
 38    away_total_xAssists                       float64
 39    away_total_key_passes                     float64
 40    away_total_xGoalsChain                    float64
 41    away_total_xGoalsBuildup                  float64
 42    away_total_yellow_cards                   int64
 43    away_total_red_cards                      int64
 44    away_total_blocked_shots                  float64
 45    away_total_saved_shots                    float64
 46    gameresult                                int64
 47    home_redCards_binary                      bool
 48    away_redCards_binary                      bool
 49    home_yellowCards_cat                      category
 50    away_yellowCards_cat                      category
 51    home_shotsOnTarget_cat                    category
 52    away_shotsOnTarget_cat                    category
 53    home_total_assists_cat                    category
```

```
54   away_total_assists_cat                       category
55   home_corners_cat                             category
56   away_corners_cat                             category
57   home_Goals_cat                               category
58   away_Goals_cat                               category
59   home_total_blocked_shots_cat                 category
60   away_total_blocked_shots_cat                 category
61   home_total_saved_shots_cat                   category
62   away_total_saved_shots_cat                   category
63   year                                         int64
64   month                                        int64
65   day                                          int64
66   time_category                                object
67   is_weekend                                   bool
68   home_cumulative_goals_for                    float64
69   away_cumulative_goals_for                    float64
70   home_cumulative_goals_against                float64
71   away_cumulative_goals_against                float64
72   home_cumulative_goal_diff                    float64
73   away_cumulative_goal_diff                    float64
74   home_avg_goals_for                           float64
75   away_avg_goals_for                           float64
76   home_avg_goals_against                       float64
77   away_avg_goals_against                       float64
78   home_goal_per_shot                           float64
79   away_goal_per_shot                           float64
80   home_goal_per_xg                             float64
81   away_goal_per_xg                             float64
82   home_rolling_goal_per_shot                   float64
83   away_rolling_goal_per_shot                   float64
84   home_rolling_goal_per_xg                     float64
85   away_rolling_goal_per_xg                     float64
86   home_blocked_shot_ratio                      float64
87   away_blocked_shot_ratio                      float64
88   home_shot_conceded_ratio                     float64
89   away_shot_conceded_ratio                     float64
90   home_discipline_score                        float64
91   away_discipline_score                        float64
92   home_discipline_score_avg                    float64
93   away_discipline_score_avg                    float64
94   home_adv_goal_diff                           float64
95   home_adv_avg_goals                           float64
96   points_diff                                  float64
97   recent_form_diff                             float64
98   interaction_goal_rate_vs_blocks              float64
99   interaction_goal_rate_vs_conceded            float64
100  interaction_xG_efficiency_vs_shots_conceded  float64
101  rolling_xG_vs_shot_conceded                  float64
102  rolling_shot_rate_vs_blocks                  float64
103  discipline_matchup_score                     float64
104  interaction_away_goal_rate_vs_home_def       float64
105  interaction_away_form_vs_home_cards          float64
106  form_vs_discipline_gap                       float64
dtypes: bool(3), category(14), datetime64[ns](1), float64(62), int64(26), object
(1)
memory usage: 9.0+ MB
```

	Variable
0	gameID
1	leagueID
2	season
3	date
4	homeTeamID
…	…
101	rolling_shot_rate_vs_blocks
102	discipline_matchup_score
103	interaction_away_goal_rate_vs_home_def
104	interaction_away_form_vs_home_cards
105	form_vs_discipline_gap

106 rows × 1 columns

Univariable Analysis

```
Index(['gameID', 'leagueID', 'season', 'date', 'homeTeamID', 'awayTeamID',
       'home_Goals', 'away_Goals', 'home_GoalsHalfTime', 'away_GoalsHalfTime',
       ...
       'interaction_goal_rate_vs_blocks', 'interaction_goal_rate_vs_conceded',
       'interaction_xG_efficiency_vs_shots_conceded',
       'rolling_xG_vs_shot_conceded', 'rolling_shot_rate_vs_blocks',
       'discipline_matchup_score', 'interaction_away_goal_rate_vs_home_def',
       'interaction_away_form_vs_home_cards', 'form_vs_discipline_gap',
       'gameresult'],
      dtype='object', length=107)
```

	gameID	leagueID	season	date	homeTeamID	awayTeamID	home_Goals	away_Goals
5437	6185	5	2014	2014-08-08 19:30:00	177	161	2	2
5438	6186	5	2014	2014-08-09 20:00:00	162	164	3	3
5439	6187	5	2014	2014-08-09 20:00:00	209	165	0	3
5440	6188	5	2014	2014-08-09 20:00:00	169	175	0	2
5441	6189	5	2014	2014-08-09 20:00:00	160	180	0	0

5 rows × 107 columns

```
Begining analysis...
Factorizing... please wait
***************************************************************************
************************
['The variable date is a date. Dates are not allowed in Table1... avoided']
***************************************************************************
*************************
***************************************************************************
*************************
***************************************************************************
*************************
['Unable to calcualte the Fisher exact test for variables away_GoalsHalfTime and g
ameresult... The p-value may be incorrect', 'Unable to calcualte the Fisher exact
test for variables home_redCards and gameresult... The p-value may be incorrect',
'Unable to calcualte the Fisher exact test for variables home_total_red_cards and
gameresult... The p-value may be incorrect', 'Unable to calcualte the Fisher exact
test for variables away_redCards and gameresult... The p-value may be incorrect',
'Unable to calcualte the Fisher exact test for variables away_total_red_cards and
gameresult... The p-value may be incorrect', 'Unable to calcualte the Fisher exact
test for variables home_Goals_cat and gameresult... The p-value may be incorrect',
'Unable to calcualte the Fisher exact test for variables away_Goals_cat and gamere
sult... The p-value may be incorrect']
------ Finished in 54.536526679992676econds -----
```

	Variables	Categories	Population	Category_1	Category_0
0	Individuals	n	12680	3172	3854
1	gameID	Mean (SD)	7,877.70 (4,753.60)	7,895.50 (4,772.90)	8,125.40 (4,827.10)
2	gameID	Median (IQR)	7,981.50 (3,812.80-12,023.20)	8,035.50 (3,868.50-12,021.00)	8,202.00 (3,986.50-12,385.80)
4	leagueID	1	2,660.00 (21.00%)	629.00 (19.80%)	845.00 (21.90%)
5	leagueID	2	2,660.00 (21.00%)	668.00 (21.10%)	837.00 (21.70%)
6	leagueID	4	2,660.00 (21.00%)	682.00 (21.50%)	764.00 (19.80%)
7	leagueID	5	2,558.00 (20.20%)	661.00 (20.80%)	757.00 (19.60%)
8	leagueID	3	2,142.00 (16.90%)	532.00 (16.80%)	651.00 (16.90%)
9	season	Mean (SD)	2,017.00 (2.00)	2,017.00 (2.00)	2,017.10 (2.00)
10	season	Median (IQR)	2,017.00 (2,015.00-2,019.00)	2,017.00 (2,015.00-2,019.00)	2,017.00 (2,015.00-2,019.00)
12	homeTeamID	Mean (SD)	135.30 (42.80)	137.20 (43.10)	138.80 (46.40)
13	homeTeamID	Median (IQR)	132.00 (101.00-163.00)	134.00 (104.00-164.00)	133.00 (101.00-167.00)
15	awayTeamID	Mean (SD)	135.30 (42.80)	135.80 (42.30)	130.50 (39.70)
16	awayTeamID	Median (IQR)	132.00 (101.00-163.00)	133.00 (102.00-163.00)	128.00 (98.00-158.00)
18	home_Goals	Mean (SD)	1.50 (1.30)	1.00 (0.80)	0.60 (0.70)
19	home_Goals	Median (IQR)	1.00 (1.00-2.00)	1.00 (0.00-1.00)	0.00 (0.00-1.00)
21	away_Goals	Mean (SD)	1.20 (1.20)	1.00 (0.80)	2.30 (1.10)
22	away_Goals	Median (IQR)	1.00 (0.00-2.00)	1.00 (0.00-1.00)	2.00 (2.00-3.00)
24	home_GoalsHalfTime	Mean (SD)	0.70 (0.80)	0.50 (0.60)	0.30 (0.50)
25	home_GoalsHalfTime	Median (IQR)	0.00 (0.00-1.00)	0.00 (0.00-1.00)	0.00 (0.00-0.00)
27	away_GoalsHalfTime	0	7,550.00 (59.50%)	1,973.00 (62.20%)	1,221.00 (31.70%)
28	away_GoalsHalfTime	1	3,847.00 (30.30%)	1,013.00 (31.90%)	1,647.00 (42.70%)
29	away_GoalsHalfTime	2	1,058.00 (8.30%)	173.00 (5.50%)	777.00 (20.20%)
30	away_GoalsHalfTime	3	193.00 (1.50%)	13.00 (0.40%)	177.00 (4.60%)

	Variables	Categories	Population	Category_1	Category_0
31	away_GoalsHalfTime	4	29.00 (0.20%)	0.00 (0.00%)	29.00 (0.80%)
32	away_GoalsHalfTime	5	3.00 (0.00%)	0.00 (0.00%)	3.00 (0.10%)
33	home_xGoals	Mean (SD)	1.40 (0.80)	1.30 (0.70)	1.00 (0.60)
34	home_xGoals	Median (IQR)	1.30 (0.80- 1.90)	1.20 (0.80- 1.70)	0.90 (0.60- 1.40)
36	home_shots	Mean (SD)	13.50 (4.90)	13.20 (4.90)	12.30 (4.80)
37	home_shots	Median (IQR)	13.00 (10.00- 17.00)	13.00 (10.00- 16.00)	12.00 (9.00- 15.00)
39	home_shotsOnTarget	Mean (SD)	4.70 (2.60)	4.10 (2.20)	3.50 (2.10)
40	home_shotsOnTarget	Median (IQR)	4.00 (3.00- 6.00)	4.00 (3.00- 5.00)	3.00 (2.00- 5.00)
42	home_deep	Mean (SD)	6.10 (3.70)	5.90 (3.60)	5.40 (3.40)
43	home_deep	Median (IQR)	6.00 (3.00- 8.00)	5.00 (3.00- 8.00)	5.00 (3.00- 7.00)
45	home_ppda	Mean (SD)	10.50 (5.60)	10.10 (5.40)	11.30 (6.70)
46	home_ppda	Median (IQR)	9.20 (6.80- 12.60)	8.90 (6.50- 12.20)	9.70 (7.00- 13.60)
48	home_fouls	Mean (SD)	12.60 (3.80)	12.90 (3.80)	12.70 (3.80)
49	home_fouls	Median (IQR)	12.60 (10.00- 15.00)	13.00 (10.00- 15.00)	12.60 (10.00- 15.00)
51	home_corners	Mean (SD)	5.40 (2.80)	5.50 (2.80)	5.20 (2.80)
52	home_corners	Median (IQR)	5.00 (3.00- 7.00)	5.00 (3.00- 7.00)	5.00 (3.00- 7.00)
54	home_yellowCards	Mean (SD)	1.90 (1.40)	2.10 (1.30)	2.10 (1.40)
55	home_yellowCards	Median (IQR)	2.00 (1.00- 3.00)	2.00 (1.00- 3.00)	2.00 (1.00- 3.00)
57	home_redCards	0	11,602.00 (91.50%)	2,878.00 (90.70%)	3,345.00 (86.80%)
58	home_redCards	1	1,016.00 (8.00%)	282.00 (8.90%)	469.00 (12.20%)
59	home_redCards	2	61.00 (0.50%)	12.00 (0.40%)	39.00 (1.00%)
60	home_redCards	3	1.00 (0.00%)	0.00 (0.00%)	1.00 (0.00%)
61	home_total_assists	Mean (SD)	1.00 (1.10)	0.60 (0.70)	0.40 (0.60)
62	home_total_assists	Median (IQR)	1.00 (0.00- 2.00)	1.00 (0.00- 1.00)	0.00 (0.00- 1.00)
64	home_total_xAssists	Mean (SD)	1.00 (0.60)	0.90 (0.50)	0.70 (0.50)
65	home_total_xAssists	Median (IQR)	0.90 (0.50- 1.40)	0.80 (0.50- 1.20)	0.60 (0.40- 1.00)

	Variables	Categories	Population	Category_1	Category_0
67	home_total_key_passes	Mean (SD)	10.00 (4.10)	9.70 (4.10)	9.00 (3.90)
68	home_total_key_passes	Median (IQR)	10.00 (7.00-13.00)	9.00 (7.00-12.00)	9.00 (6.00-11.00)
70	home_total_xGoalsChain	Mean (SD)	4.00 (3.30)	3.30 (2.40)	2.50 (2.00)
71	home_total_xGoalsChain	Median (IQR)	3.10 (1.60-5.40)	2.70 (1.40-4.50)	2.00 (1.10-3.50)
73	home_total_xGoalsBuildup	Mean (SD)	2.30 (2.30)	1.90 (1.70)	1.50 (1.40)
74	home_total_xGoalsBuildup	Median (IQR)	1.60 (0.70-3.20)	1.40 (0.70-2.70)	1.00 (0.50-2.00)
76	home_total_yellow_cards	Mean (SD)	1.90 (1.30)	2.00 (1.30)	2.00 (1.30)
77	home_total_yellow_cards	Median (IQR)	2.00 (1.00-3.00)	2.00 (1.00-3.00)	2.00 (1.00-3.00)
79	home_total_red_cards	0	11,616.00 (91.60%)	2,880.00 (90.80%)	3,352.00 (87.00%)
80	home_total_red_cards	1	1,007.00 (7.90%)	281.00 (8.90%)	465.00 (12.10%)
81	home_total_red_cards	2	56.00 (0.40%)	11.00 (0.30%)	36.00 (0.90%)
82	home_total_red_cards	3	1.00 (0.00%)	0.00 (0.00%)	1.00 (0.00%)
83	home_total_blocked_shots	Mean (SD)	3.50 (2.30)	3.60 (2.40)	3.50 (2.30)
84	home_total_blocked_shots	Median (IQR)	3.00 (2.00-5.00)	3.00 (2.00-5.00)	3.00 (2.00-5.00)
86	home_total_saved_shots	Mean (SD)	3.30 (2.10)	3.20 (2.10)	3.00 (2.00)
87	home_total_saved_shots	Median (IQR)	3.00 (2.00-4.00)	3.00 (2.00-4.00)	3.00 (2.00-4.00)
89	away_xGoals	Mean (SD)	1.20 (0.80)	1.10 (0.60)	1.70 (0.90)
90	away_xGoals	Median (IQR)	1.00 (0.60-1.60)	1.00 (0.60-1.50)	1.60 (1.10-2.20)
92	away_shots	Mean (SD)	11.00 (4.30)	10.80 (4.30)	12.40 (4.30)
93	away_shots	Median (IQR)	11.00 (8.00-14.00)	10.00 (8.00- 14.00)	12.00 (9.00- 15.00)
95	away_shotsOnTarget	Mean (SD)	3.90 (2.30)	3.60 (2.00)	5.30 (2.30)
96	away_shotsOnTarget	Median (IQR)	4.00 (2.00-5.00)	3.00 (2.00-5.00)	5.00 (4.00-7.00)
98	away_deep	Mean (SD)	4.80 (3.00)	4.70 (3.00)	5.70 (3.10)
99	away_deep	Median (IQR)	4.00 (3.00-7.00)	4.00 (2.00-6.00)	5.00 (3.00-8.00)
101	away_ppda	Mean (SD)	11.80 (6.90)	11.20 (6.00)	11.10 (5.60)
102	away_ppda	Median (IQR)	10.20 (7.50-14.20)	10.00 (7.30- 13.60)	9.90 (7.40-13.20)

	Variables	Categories	Population	Category_1	Category_0
104	away_fouls	Mean (SD)	13.00 (4.10)	13.50 (4.10)	12.80 (4.10)
105	away_fouls	Median (IQR)	13.00 (10.00-16.00)	13.00 (11.00-16.00)	13.00 (10.00-15.00)
107	away_corners	Mean (SD)	4.50 (2.60)	4.40 (2.60)	4.60 (2.60)
108	away_corners	Median (IQR)	4.00 (3.00-6.00)	4.00 (2.00-6.00)	4.00 (3.00-6.00)
110	away_yellowCards	Mean (SD)	2.10 (1.40)	2.30 (1.40)	2.10 (1.40)
111	away_yellowCards	Median (IQR)	2.00 (1.00-3.00)	2.00 (1.00-3.00)	2.00 (1.00-3.00)
113	away_redCards	0	11,284.00 (89.00%)	2,828.00 (89.20%)	3,605.00 (93.50%)
114	away_redCards	1	1,317.00 (10.40%)	331.00 (10.40%)	241.00 (6.30%)
115	away_redCards	2	78.00 (0.60%)	13.00 (0.40%)	8.00 (0.20%)
116	away_redCards	3	1.00 (0.00%)	0.00 (0.00%)	0.00 (0.00%)
117	away_total_assists	Mean (SD)	0.80 (1.00)	0.70 (0.70)	1.60 (1.10)
118	away_total_assists	Median (IQR)	1.00 (0.00-1.00)	1.00 (0.00-1.00)	1.00 (1.00-2.00)
120	away_total_xAssists	Mean (SD)	0.80 (0.60)	0.80 (0.50)	1.20 (0.70)
121	away_total_xAssists	Median (IQR)	0.70 (0.40-1.10)	0.70 (0.40-1.00)	1.10 (0.70-1.70)
123	away_total_key_passes	Mean (SD)	8.20 (3.60)	8.00 (3.50)	9.40 (3.60)
124	away_total_key_passes	Median (IQR)	8.00 (6.00-10.00)	8.00 (5.00-10.00)	9.00 (7.00-12.00)
126	away_total_xGoalsChain	Mean (SD)	3.20 (2.80)	2.80 (2.30)	5.00 (3.50)
127	away_total_xGoalsChain	Median (IQR)	2.40 (1.20-4.40)	2.20 (1.10-3.90)	4.20 (2.40-6.60)
129	away_total_xGoalsBuildup	Mean (SD)	1.90 (2.00)	1.60 (1.70)	3.00 (2.60)
130	away_total_xGoalsBuildup	Median (IQR)	1.20 (0.50-2.50)	1.10 (0.50-2.20)	2.30 (1.10-4.10)
132	away_total_yellow_cards	Mean (SD)	2.10 (1.30)	2.20 (1.30)	2.00 (1.30)
133	away_total_yellow_cards	Median (IQR)	2.00 (1.00-3.00)	2.00 (1.00-3.00)	2.00 (1.00-3.00)
135	away_total_red_cards	0	11,298.00 (89.10%)	2,835.00 (89.40%)	3,609.00 (93.60%)
136	away_total_red_cards	1	1,306.00 (10.30%)	324.00 (10.20%)	237.00 (6.10%)
137	away_total_red_cards	2	75.00 (0.60%)	13.00 (0.40%)	8.00 (0.20%)
138	away_total_red_cards	3	1.00 (0.00%)	0.00 (0.00%)	0.00 (0.00%)
139	away_total_blocked_shots	Mean (SD)	2.80 (2.10)	2.90 (2.10)	2.90 (2.10)

	Variables	Categories	Population	Category_1	Category_0
140	away_total_blocked_shots	Median (IQR)	2.00 (1.00-4.00)	2.00 (1.00-4.00)	3.00 (1.00-4.00)
142	away_total_saved_shots	Mean (SD)	2.70 (1.90)	2.60 (1.80)	3.00 (2.00)
143	away_total_saved_shots	Median (IQR)	2.00 (1.00-4.00)	2.00 (1.00-4.00)	3.00 (2.00-4.00)
145	home_redCards_binary	False	11,602.00 (91.50%)	2,878.00 (90.70%)	3,345.00 (86.80%)
146	home_redCards_binary	True	1,078.00 (8.50%)	294.00 (9.30%)	509.00 (13.20%)
147	away_redCards_binary	False	11,284.00 (89.00%)	2,828.00 (89.20%)	3,605.00 (93.50%)
148	away_redCards_binary	True	1,396.00 (11.00%)	344.00 (10.80%)	249.00 (6.50%)
149	home_yellowCards_cat	2	3,631.00 (28.60%)	949.00 (29.90%)	1,161.00 (30.10%)
150	home_yellowCards_cat	1	3,433.00 (27.10%)	775.00 (24.40%)	977.00 (25.40%)
151	home_yellowCards_cat	3	2,252.00 (17.80%)	622.00 (19.60%)	731.00 (19.00%)
152	home_yellowCards_cat	0	1,813.00 (14.30%)	372.00 (11.70%)	438.00 (11.40%)
153	home_yellowCards_cat	4	1,038.00 (8.20%)	320.00 (10.10%)	361.00 (9.40%)
154	home_yellowCards_cat	5+	513.00 (4.00%)	134.00 (4.20%)	186.00 (4.80%)
155	away_yellowCards_cat	2	3,605.00 (28.40%)	910.00 (28.70%)	1,050.00 (27.20%)
156	away_yellowCards_cat	1	3,045.00 (24.00%)	707.00 (22.30%)	969.00 (25.10%)
157	away_yellowCards_cat	3	2,691.00 (21.20%)	688.00 (21.70%)	798.00 (20.70%)
158	away_yellowCards_cat	0	1,356.00 (10.70%)	290.00 (9.10%)	472.00 (12.20%)
159	away_yellowCards_cat	4	1,319.00 (10.40%)	392.00 (12.40%)	359.00 (9.30%)
160	away_yellowCards_cat	5+	664.00 (5.20%)	185.00 (5.80%)	206.00 (5.30%)
161	home_shotsOnTarget_cat	4	2,033.00 (16.00%)	557.00 (17.60%)	631.00 (16.40%)
162	home_shotsOnTarget_cat	3	2,011.00 (15.90%)	590.00 (18.60%)	777.00 (20.20%)
163	home_shotsOnTarget_cat	5	1,860.00 (14.70%)	473.00 (14.90%)	491.00 (12.70%)
164	home_shotsOnTarget_cat	6	1,477.00 (11.60%)	347.00 (10.90%)	301.00 (7.80%)
165	home_shotsOnTarget_cat	2	1,469.00 (11.60%)	457.00 (14.40%)	719.00 (18.70%)

	Variables	Categories	Population	Category_1	Category_0
166	home_shotsOnTarget_cat	7	1,071.00 (8.40%)	NaN	NaN
167	home_shotsOnTarget_cat	Other	2,003.00 (15.80%)	291.00 (9.20%)	331.00 (8.60%)
168	away_shotsOnTarget_cat	3	2,333.00 (18.40%)	646.00 (20.40%)	525.00 (13.60%)
169	away_shotsOnTarget_cat	4	2,114.00 (16.70%)	575.00 (18.10%)	664.00 (17.20%)
170	away_shotsOnTarget_cat	2	2,081.00 (16.40%)	597.00 (18.80%)	273.00 (7.10%)
171	away_shotsOnTarget_cat	5	1,627.00 (12.80%)	406.00 (12.80%)	661.00 (17.20%)
172	away_shotsOnTarget_cat	1	1,277.00 (10.10%)	320.00 (10.10%)	NaN
173	away_shotsOnTarget_cat	6	1,191.00 (9.40%)	266.00 (8.40%)	602.00 (15.60%)
174	away_shotsOnTarget_cat	Other	1,319.00 (10.40%)	224.00 (7.10%)	449.00 (11.70%)
175	home_total_assists_cat	0	4,785.00 (37.70%)	1,548.00 (48.80%)	2,564.00 (66.50%)
176	home_total_assists_cat	1	4,382.00 (34.60%)	1,249.00 (39.40%)	1,099.00 (28.50%)
177	home_total_assists_cat	2	2,288.00 (18.00%)	330.00 (10.40%)	178.00 (4.60%)
178	home_total_assists_cat	3+	1,225.00 (9.70%)	45.00 (1.40%)	13.00 (0.30%)
179	away_total_assists_cat	0	5,812.00 (45.80%)	1,492.00 (47.00%)	448.00 (11.60%)
180	away_total_assists_cat	1	4,335.00 (34.20%)	1,290.00 (40.70%)	1,498.00 (38.90%)
181	away_total_assists_cat	2	1,743.00 (13.70%)	348.00 (11.00%)	1,177.00 (30.50%)
182	away_total_assists_cat	3+	790.00 (6.20%)	42.00 (1.30%)	731.00 (19.00%)
183	home_corners_cat	4	1,778.00 (14.00%)	439.00 (13.80%)	505.00 (13.10%)
184	home_corners_cat	5	1,758.00 (13.90%)	418.00 (13.20%)	526.00 (13.60%)
185	home_corners_cat	3	1,544.00 (12.20%)	376.00 (11.90%)	477.00 (12.40%)
186	home_corners_cat	6	1,533.00 (12.10%)	384.00 (12.10%)	462.00 (12.00%)
187	home_corners_cat	7	1,265.00 (10.00%)	328.00 (10.30%)	358.00 (9.30%)
188	home_corners_cat	2	1,118.00 (8.80%)	251.00 (7.90%)	384.00 (10.00%)
189	home_corners_cat	Other	2,701.00 (21.30%)	731.00 (23.00%)	844.00 (21.90%)

	Variables	Categories	Population	Category_1	Category_0
190	away_corners_cat	4	1,987.00 (15.70%)	472.00 (14.90%)	615.00 (16.00%)
191	away_corners_cat	3	1,952.00 (15.40%)	493.00 (15.50%)	604.00 (15.70%)
192	away_corners_cat	5	1,723.00 (13.60%)	422.00 (13.30%)	542.00 (14.10%)
193	away_corners_cat	2	1,595.00 (12.60%)	447.00 (14.10%)	436.00 (11.30%)
194	away_corners_cat	6	1,345.00 (10.60%)	333.00 (10.50%)	418.00 (10.80%)
195	away_corners_cat	1	1,121.00 (8.80%)	279.00 (8.80%)	NaN
196	away_corners_cat	Other	1,941.00 (15.30%)	479.00 (15.10%)	589.00 (15.30%)
197	home_Goals_cat	1	4,086.00 (32.20%)	1,478.00 (46.60%)	1,427.00 (37.00%)
198	home_Goals_cat	2	3,099.00 (24.40%)	637.00 (20.10%)	352.00 (9.10%)
199	home_Goals_cat	0	2,943.00 (23.20%)	915.00 (28.80%)	2,028.00 (52.60%)
200	home_Goals_cat	3	1,571.00 (12.40%)	132.00 (4.20%)	42.00 (1.10%)
201	home_Goals_cat	4+	981.00 (7.70%)	10.00 (0.30%)	5.00 (0.10%)
202	away_Goals_cat	1	4,458.00 (35.20%)	1,478.00 (46.60%)	956.00 (24.80%)
203	away_Goals_cat	0	4,031.00 (31.80%)	915.00 (28.80%)	0.00 (0.00%)
204	away_Goals_cat	2	2,522.00 (19.90%)	637.00 (20.10%)	1,434.00 (37.20%)
205	away_Goals_cat	3	1,108.00 (8.70%)	132.00 (4.20%)	918.00 (23.80%)
206	away_Goals_cat	4+	561.00 (4.40%)	10.00 (0.30%)	546.00 (14.20%)
207	home_total_blocked_shots_cat	3	2,341.00 (18.50%)	565.00 (17.80%)	719.00 (18.70%)
208	home_total_blocked_shots_cat	2	2,318.00 (18.30%)	582.00 (18.30%)	685.00 (17.80%)
209	home_total_blocked_shots_cat	4	1,810.00 (14.30%)	453.00 (14.30%)	556.00 (14.40%)
210	home_total_blocked_shots_cat	1	1,766.00 (13.90%)	407.00 (12.80%)	548.00 (14.20%)
211	home_total_blocked_shots_cat	5	1,382.00 (10.90%)	346.00 (10.90%)	421.00 (10.90%)
212	home_total_blocked_shots_cat	6	927.00 (7.30%)	223.00 (7.00%)	300.00 (7.80%)
213	home_total_blocked_shots_cat	Other	1,315.00 (10.40%)	391.00 (12.30%)	383.00 (9.90%)

	Variables	Categories	Population	Category_1	Category_0
214	away_total_blocked_shots_cat	2	2,727.00 (21.50%)	654.00 (20.60%)	802.00 (20.80%)
215	away_total_blocked_shots_cat	1	2,344.00 (18.50%)	600.00 (18.90%)	683.00 (17.70%)
216	away_total_blocked_shots_cat	3	2,244.00 (17.70%)	550.00 (17.30%)	699.00 (18.10%)
217	away_total_blocked_shots_cat	4	1,649.00 (13.00%)	410.00 (12.90%)	529.00 (13.70%)
218	away_total_blocked_shots_cat	0	1,311.00 (10.30%)	336.00 (10.60%)	390.00 (10.10%)
219	away_total_blocked_shots_cat	5	1,046.00 (8.20%)	285.00 (9.00%)	305.00 (7.90%)
220	away_total_blocked_shots_cat	Other	741.00 (5.80%)	190.00 (6.00%)	256.00 (6.60%)
221	home_total_saved_shots_cat	3	2,553.00 (20.10%)	636.00 (20.10%)	799.00 (20.70%)
222	home_total_saved_shots_cat	2	2,515.00 (19.80%)	645.00 (20.30%)	817.00 (21.20%)
223	home_total_saved_shots_cat	4	1,877.00 (14.80%)	460.00 (14.50%)	580.00 (15.00%)
224	home_total_saved_shots_cat	1	1,857.00 (14.60%)	476.00 (15.00%)	618.00 (16.00%)
225	home_total_saved_shots_cat	5	1,264.00 (10.00%)	326.00 (10.30%)	307.00 (8.00%)
226	home_total_saved_shots_cat	6	848.00 (6.70%)	NaN	NaN
227	home_total_saved_shots_cat	Other	968.00 (7.60%)	219.00 (6.90%)	215.00 (5.60%)
228	away_total_saved_shots_cat	2	2,861.00 (22.60%)	748.00 (23.60%)	805.00 (20.90%)
229	away_total_saved_shots_cat	3	2,473.00 (19.50%)	606.00 (19.10%)	774.00 (20.10%)
230	away_total_saved_shots_cat	1	2,453.00 (19.30%)	648.00 (20.40%)	610.00 (15.80%)
231	away_total_saved_shots_cat	4	1,655.00 (13.10%)	387.00 (12.20%)	573.00 (14.90%)
232	away_total_saved_shots_cat	0	1,172.00 (9.20%)	292.00 (9.20%)	288.00 (7.50%)
233	away_total_saved_shots_cat	5	988.00 (7.80%)	252.00 (7.90%)	360.00 (9.30%)
234	away_total_saved_shots_cat	Other	512.00 (4.00%)	114.00 (3.60%)	227.00 (5.90%)
235	year	Mean (SD)	2,017.50 (2.10)	2,017.50 (2.10)	2,017.60 (2.10)
236	year	Median (IQR)	2,017.00 (2,016.00-2,019.00)	2,017.00 (2,016.00-2,019.00)	2,018.00 (2,016.00-2,019.00)
238	month	Mean (SD)	6.50 (3.80)	6.60 (3.90)	6.50 (3.80)

	Variables	Categories	Population	Category_1	Category_0
239	month	Median (IQR)	6.00 (3.00- 10.00)	7.00 (3.00- 10.00)	6.00 (3.00- 10.00)
241	day	Mean (SD)	16.00 (8.80)	16.10 (8.80)	16.10 (8.70)
242	day	Median (IQR)	17.00 (8.00- 23.00)	17.00 (8.00- 23.00)	17.00 (9.00- 23.00)
244	time_category	Night	6,169.00 (48.70%)	1,551.00 (48.90%)	1,799.00 (46.70%)
245	time_category	After Noon	5,278.00 (41.60%)	1,307.00 (41.20%)	1,659.00 (43.00%)
246	time_category	Noon	1,233.00 (9.70%)	314.00 (9.90%)	396.00 (10.30%)
247	is_weekend	True	10,632.00 (83.80%)	2,658.00 (83.80%)	3,238.00 (84.00%)
248	is_weekend	False	2,048.00 (16.20%)	514.00 (16.20%)	616.00 (16.00%)
249	home_cumulative_goals_for	Mean (SD)	85.90 (70.30)	87.10 (71.80)	85.10 (69.90)
250	home_cumulative_goals_for	Median (IQR)	69.00 (29.00- 129.00)	71.00 (30.00- 131.00)	69.00 (28.00- 126.80)
252	away_cumulative_goals_for	Mean (SD)	66.10 (55.00)	67.00 (56.30)	65.30 (54.30)
253	away_cumulative_goals_for	Median (IQR)	52.00 (22.00- 98.00)	52.00 (21.00- 100.00)	52.00 (21.00- 96.00)
255	home_cumulative_goals_against	Mean (SD)	59.60 (41.70)	60.00 (42.20)	59.20 (41.50)
256	home_cumulative_goals_against	Median (IQR)	54.00 (24.00- 90.00)	54.00 (24.00- 90.00)	54.00 (24.00- 89.00)
258	away_cumulative_goals_against	Mean (SD)	78.00 (53.40)	78.10 (54.20)	77.50 (52.90)
259	away_cumulative_goals_against	Median (IQR)	70.00 (32.00- 119.00)	71.00 (31.00- 118.00)	71.00 (32.00- 118.00)
261	home_cumulative_goal_diff	Mean (SD)	26.20 (48.70)	27.10 (50.20)	25.90 (48.80)
262	home_cumulative_goal_diff	Median (IQR)	8.00 (-3.00- 39.00)	8.00 (-2.00- 40.00)	8.00 (-3.00- 38.00)
264	away_cumulative_goal_diff	Mean (SD)	-11.90 (38.10)	-11.10 (39.10)	-12.10 (37.90)
265	away_cumulative_goal_diff	Median (IQR)	-13.00 (-34.00- 0.00)	-13.00 (-34.00- 0.00)	-13.00 (-35.00- 0.00)
267	home_avg_goals_for	Mean (SD)	1.50 (0.60)	1.50 (0.60)	1.50 (0.60)
268	home_avg_goals_for	Median (IQR)	1.40 (1.20- 1.80)	1.40 (1.20- 1.80)	1.40 (1.20- 1.80)

	Variables	Categories	Population	Category_1	Category_0
270	away_avg_goals_for	Mean (SD)	1.20 (0.40)	1.20 (0.50)	1.20 (0.40)
271	away_avg_goals_for	Median (IQR)	1.10 (0.90-1.40)	1.10 (0.90-1.40)	1.10 (0.90-1.40)
273	home_avg_goals_against	Mean (SD)	1.10 (0.40)	1.10 (0.40)	1.10 (0.40)
274	home_avg_goals_against	Median (IQR)	1.10 (0.90-1.30)	1.10 (0.90-1.30)	1.10 (0.90-1.30)
276	away_avg_goals_against	Mean (SD)	1.50 (0.40)	1.50 (0.40)	1.50 (0.40)
277	away_avg_goals_against	Median (IQR)	1.50 (1.30-1.70)	1.50 (1.20-1.70)	1.50 (1.20-1.70)
279	home_goal_per_shot	Mean (SD)	0.10 (0.00)	0.10 (0.00)	0.10 (0.00)
280	home_goal_per_shot	Median (IQR)	0.10 (0.10-0.10)	0.10 (0.10-0.10)	0.10 (0.10-0.10)
282	away_goal_per_shot	Mean (SD)	0.10 (0.00)	0.10 (0.00)	0.10 (0.00)
283	away_goal_per_shot	Median (IQR)	0.10 (0.10-0.10)	0.10 (0.10-0.10)	0.10 (0.10-0.10)
285	home_goal_per_xg	Mean (SD)	1.00 (0.30)	1.10 (0.30)	1.10 (0.30)
286	home_goal_per_xg	Median (IQR)	1.10 (1.00-1.10)	1.10 (1.00-1.10)	1.10 (1.00-1.20)
288	away_goal_per_xg	Mean (SD)	1.00 (0.40)	1.00 (0.60)	1.00 (0.30)
289	away_goal_per_xg	Median (IQR)	1.00 (0.90-1.10)	1.00 (0.90-1.10)	1.00 (0.90-1.10)
291	home_rolling_goal_per_shot	Mean (SD)	0.10 (0.10)	0.10 (0.10)	0.10 (0.10)
292	home_rolling_goal_per_shot	Median (IQR)	0.10 (0.10-0.10)	0.10 (0.10-0.10)	0.10 (0.10-0.10)
294	away_rolling_goal_per_shot	Mean (SD)	0.10 (0.10)	0.10 (0.10)	0.10 (0.10)
295	away_rolling_goal_per_shot	Median (IQR)	0.10 (0.10-0.10)	0.10 (0.10-0.10)	0.10 (0.10-0.10)
297	home_rolling_goal_per_xg	Mean (SD)	1.00 (0.40)	1.00 (0.40)	1.00 (0.40)
298	home_rolling_goal_per_xg	Median (IQR)	1.00 (0.80-1.30)	1.00 (0.80-1.30)	1.00 (0.80-1.30)
300	away_rolling_goal_per_xg	Mean (SD)	1.00 (0.40)	1.00 (0.40)	1.00 (0.40)
301	away_rolling_goal_per_xg	Median (IQR)	1.00 (0.70-1.20)	1.00 (0.70-1.20)	1.00 (0.70-1.20)
303	home_blocked_shot_ratio	Mean (SD)	0.40 (0.50)	0.40 (0.50)	0.40 (0.40)
304	home_blocked_shot_ratio	Median (IQR)	0.30 (0.10-0.50)	0.30 (0.10-0.50)	0.30 (0.10-0.50)
306	away_blocked_shot_ratio	Mean (SD)	0.30 (0.30)	0.30 (0.30)	0.30 (0.30)
307	away_blocked_shot_ratio	Median (IQR)	0.20 (0.10-0.30)	0.20 (0.10-0.30)	0.20 (0.10-0.30)
309	home_shot_conceded_ratio	Mean (SD)	0.40 (0.20)	0.40 (0.20)	0.30 (0.20)
310	home_shot_conceded_ratio	Median (IQR)	0.30 (0.20-0.50)	0.30 (0.20-0.50)	0.30 (0.20-0.50)
312	away_shot_conceded_ratio	Mean (SD)	0.40 (0.20)	0.30 (0.20)	0.40 (0.20)

	Variables	Categories	Population	Category_1	Category_0
313	away_shot_conceded_ratio	Median (IQR)	0.30 (0.20-0.40)	0.30 (0.20-0.50)	0.30 (0.20-0.40)
315	home_discipline_score	Mean (SD)	-1,020.60 (696.80)	-1,031.70 (710.50)	-1,008.90 (688.50)
316	home_discipline_score	Median (IQR)	-934.00 (-1,529.70--423.80)	-953.50 (-1,558.30--415.00)	-932.80 (-1,512.60--414.60)
318	away_discipline_score	Mean (SD)	-1,113.30 (760.00)	-1,122.30 (774.90)	-1,100.30 (749.60)
319	away_discipline_score	Median (IQR)	-1,017.50 (-1,671.00--463.80)	-1,037.50 (-1,697.20--455.80)	-1,012.00 (-1,652.00--454.20)
321	home_discipline_score_avg	Mean (SD)	-19.20 (3.80)	-19.20 (3.70)	-19.10 (3.90)
322	home_discipline_score_avg	Median (IQR)	-19.50 (-21.60--17.00)	-19.40 (-21.40--16.90)	-19.40 (-21.50--17.00)
324	away_discipline_score_avg	Mean (SD)	-21.00 (4.00)	-20.90 (4.00)	-20.90 (4.10)
325	away_discipline_score_avg	Median (IQR)	-21.30 (-23.50--18.40)	-21.20 (-23.50--18.30)	-21.20 (-23.40--18.30)
327	home_adv_goal_diff	Mean (SD)	38.10 (30.40)	38.10 (31.00)	38.10 (30.60)
328	home_adv_goal_diff	Median (IQR)	33.00 (14.00-56.00)	33.00 (13.00-56.00)	34.00 (14.00-56.00)
330	home_adv_avg_goals	Mean (SD)	0.30 (0.40)	0.40 (0.40)	0.40 (0.40)
331	home_adv_avg_goals	Median (IQR)	0.40 (0.20-0.50)	0.40 (0.20-0.50)	0.40 (0.20-0.50)
333	points_diff	Mean (SD)	25.30 (23.70)	25.50 (24.70)	25.10 (24.00)
334	points_diff	Median (IQR)	22.00 (8.00-38.00)	21.00 (8.00- 38.00)	21.00 (8.00- 37.00)
336	recent_form_diff	Mean (SD)	2.10 (3.90)	2.00 (3.90)	2.10 (3.90)
337	recent_form_diff	Median (IQR)	2.00 (0.00-5.00)	2.00 (0.00-5.00)	2.00 (0.00-5.00)
339	interaction_goal_rate_vs_blocks	Mean (SD)	1,129.60 (3,424.00)	1,209.20 (3,537.90)	1,141.00 (3,434.80)
340	interaction_goal_rate_vs_blocks	Median (IQR)	0.60 (0.30-1.20)	0.60 (0.30-1.20)	0.60 (0.30-1.20)
342	interaction_goal_rate_vs_conceded	Mean (SD)	70.90 (1,044.10)	79.90 (1,121.90)	60.30 (961.80)
343	interaction_goal_rate_vs_conceded	Median (IQR)	0.10 (0.10-0.10)	0.10 (0.10-0.10)	0.10 (0.10-0.10)
345	interaction_xG_efficiency_vs_shots_conceded	Mean (SD)	683.50 (10,036.60)	717.60 (9,617.00)	695.70 (11,761.70)
346	interaction_xG_efficiency_vs_shots_conceded	Median (IQR)	0.70 (0.60-0.90)	0.70 (0.60-0.90)	0.70 (0.60-0.90)

	Variables	Categories	Population	Category_1	Category_0
348	rolling_xG_vs_shot_conceded	Mean (SD)	2,125.60 (15,770.80)	2,397.20 (16,653.80)	1,901.60 (14,932.40)
349	rolling_xG_vs_shot_conceded	Median (IQR)	2.90 (1.90-4.50)	2.80 (1.90-4.50)	2.80 (1.90-4.40)
351	rolling_shot_rate_vs_blocks	Mean (SD)	1,098.00 (3,611.30)	1,168.60 (3,739.30)	1,126.40 (3,633.10)
352	rolling_shot_rate_vs_blocks	Median (IQR)	0.50 (0.20-1.20)	0.50 (0.20-1.20)	0.50 (0.30-1.20)
354	discipline_matchup_score	Mean (SD)	29.10 (553.40)	36.00 (620.90)	27.10 (551.70)
355	discipline_matchup_score	Median (IQR)	0.00 (0.00-0.00)	0.00 (0.00-0.00)	0.00 (0.00-0.00)
357	interaction_away_goal_rate_vs_home_def	Mean (SD)	690.90 (2,740.10)	686.80 (2,698.70)	700.90 (2,727.20)
358	interaction_away_goal_rate_vs_home_def	Median (IQR)	0.30 (0.20-0.70)	0.40 (0.20-0.70)	0.30 (0.20-0.70)
360	interaction_away_form_vs_home_cards	Mean (SD)	209.10 (5,399.10)	108.50 (3,049.60)	183.40 (4,960.10)
361	interaction_away_form_vs_home_cards	Median (IQR)	0.10 (0.00-0.10)	0.10 (0.00-0.10)	0.10 (0.00-0.10)
363	form_vs_discipline_gap	Mean (SD)	-2.40 (297.40)	2.20 (198.00)	-8.90 (384.20)
364	form_vs_discipline_gap	Median (IQR)	0.00 (-0.00- 0.00)	0.00 (-0.00-0.00)	0.00 (-0.00-0.00)

	Variables	Categories	Population	Category_1	Category_0	Category_2
1	gameID	Mean (SD)	7,877.70 (4,753.60)	7,895.50 (4,772.90)	8,125.40 (4,827.10)	7,698.90 (4,683.90)
9	season	Mean (SD)	2,017.00 (2.00)	2,017.00 (2.00)	2,017.10 (2.00)	2,016.90 (2.00)
12	homeTeamID	Mean (SD)	135.30 (42.80)	137.20 (43.10)	138.80 (46.40)	131.90 (39.60)
15	awayTeamID	Mean (SD)	135.30 (42.80)	135.80 (42.30)	130.50 (39.70)	138.40 (44.70)
18	home_Goals	Mean (SD)	1.50 (1.30)	1.00 (0.80)	0.60 (0.70)	2.50 (1.20)
21	away_Goals	Mean (SD)	1.20 (1.20)	1.00 (0.80)	2.30 (1.10)	0.60 (0.70)
24	home_GoalsHalfTime	Mean (SD)	0.70 (0.80)	0.50 (0.60)	0.30 (0.50)	1.10 (0.90)
28	away_GoalsHalfTime	1	3,847.00 (30.30%)	1,013.00 (31.90%)	1,647.00 (42.70%)	1,187.00 (21.00%)
33	home_xGoals	Mean (SD)	1.40 (0.80)	1.30 (0.70)	1.00 (0.60)	1.80 (0.80)
36	home_shots	Mean (SD)	13.50 (4.90)	13.20 (4.90)	12.30 (4.80)	14.40 (4.80)
39	home_shotsOnTarget	Mean (SD)	4.70 (2.60)	4.10 (2.20)	3.50 (2.10)	5.90 (2.60)
42	home_deep	Mean (SD)	6.10 (3.70)	5.90 (3.60)	5.40 (3.40)	6.80 (3.80)
45	home_ppda	Mean (SD)	10.50 (5.60)	10.10 (5.40)	11.30 (6.70)	10.20 (4.80)
48	home_fouls	Mean (SD)	12.60 (3.80)	12.90 (3.80)	12.70 (3.80)	12.40 (3.90)
51	home_corners	Mean (SD)	5.40 (2.80)	5.50 (2.80)	5.20 (2.80)	5.40 (2.70)
54	home_yellowCards	Mean (SD)	1.90 (1.40)	2.10 (1.30)	2.10 (1.40)	1.70 (1.30)
58	home_redCards	1	1,016.00 (8.00%)	282.00 (8.90%)	469.00 (12.20%)	265.00 (4.70%)
61	home_total_assists	Mean (SD)	1.00 (1.10)	0.60 (0.70)	0.40 (0.60)	1.70 (1.10)
64	home_total_xAssists	Mean (SD)	1.00 (0.60)	0.90 (0.50)	0.70 (0.50)	1.20 (0.60)
67	home_total_key_passes	Mean (SD)	10.00 (4.10)	9.70 (4.10)	9.00 (3.90)	10.90 (4.10)
70	home_total_xGoalsChain	Mean (SD)	4.00 (3.30)	3.30 (2.40)	2.50 (2.00)	5.40 (3.80)
73	home_total_xGoalsBuildup	Mean (SD)	2.30 (2.30)	1.90 (1.70)	1.50 (1.40)	3.20 (2.80)
76	home_total_yellow_cards	Mean (SD)	1.90 (1.30)	2.00 (1.30)	2.00 (1.30)	1.70 (1.30)
80	home_total_red_cards	1	1,007.00 (7.90%)	281.00 (8.90%)	465.00 (12.10%)	261.00 (4.60%)
83	home_total_blocked_shots	Mean (SD)	3.50 (2.30)	3.60 (2.40)	3.50 (2.30)	3.40 (2.30)
86	home_total_saved_shots	Mean (SD)	3.30 (2.10)	3.20 (2.10)	3.00 (2.00)	3.50 (2.20)
89	away_xGoals	Mean (SD)	1.20 (0.80)	1.10 (0.60)	1.70 (0.90)	0.90 (0.60)
92	away_shots	Mean (SD)	11.00 (4.30)	10.80 (4.30)	12.40 (4.30)	10.20 (4.20)
95	away_shotsOnTarget	Mean (SD)	3.90 (2.30)	3.60 (2.00)	5.30 (2.30)	3.10 (1.90)
98	away_deep	Mean (SD)	4.80 (3.00)	4.70 (3.00)	5.70 (3.10)	4.30 (2.80)

	Variables	Categories	Population	Category_1	Category_0	Category_2
101	away_ppda	Mean (SD)	11.80 (6.90)	11.20 (6.00)	11.10 (5.60)	12.70 (8.10)
104	away_fouls	Mean (SD)	13.00 (4.10)	13.50 (4.10)	12.80 (4.10)	12.90 (4.10)
107	away_corners	Mean (SD)	4.50 (2.60)	4.40 (2.60)	4.60 (2.60)	4.40 (2.60)
110	away_yellowCards	Mean (SD)	2.10 (1.40)	2.30 (1.40)	2.10 (1.40)	2.10 (1.40)
114	away_redCards	1	1,317.00 (10.40%)	331.00 (10.40%)	241.00 (6.30%)	745.00 (13.20%)
117	away_total_assists	Mean (SD)	0.80 (1.00)	0.70 (0.70)	1.60 (1.10)	0.40 (0.60)
120	away_total_xAssists	Mean (SD)	0.80 (0.60)	0.80 (0.50)	1.20 (0.70)	0.60 (0.40)
123	away_total_key_passes	Mean (SD)	8.20 (3.60)	8.00 (3.50)	9.40 (3.60)	7.50 (3.40)
126	away_total_xGoalsChain	Mean (SD)	3.20 (2.80)	2.80 (2.30)	5.00 (3.50)	2.20 (1.80)
129	away_total_xGoalsBuildup	Mean (SD)	1.90 (2.00)	1.60 (1.70)	3.00 (2.60)	1.30 (1.30)
132	away_total_yellow_cards	Mean (SD)	2.10 (1.30)	2.20 (1.30)	2.00 (1.30)	2.10 (1.30)
136	away_total_red_cards	1	1,306.00 (10.30%)	324.00 (10.20%)	237.00 (6.10%)	745.00 (13.20%)
139	away_total_blocked_shots	Mean (SD)	2.80 (2.10)	2.90 (2.10)	2.90 (2.10)	2.80 (2.00)
142	away_total_saved_shots	Mean (SD)	2.70 (1.90)	2.60 (1.80)	3.00 (2.00)	2.50 (1.80)
146	home_redCards_binary	True	1,078.00 (8.50%)	294.00 (9.30%)	509.00 (13.20%)	275.00 (4.90%)
148	away_redCards_binary	True	1,396.00 (11.00%)	344.00 (10.80%)	249.00 (6.50%)	803.00 (14.20%)
150	home_yellowCards_cat	1	3,433.00 (27.10%)	775.00 (24.40%)	977.00 (25.40%)	1,681.00 (29.70%)
156	away_yellowCards_cat	1	3,045.00 (24.00%)	707.00 (22.30%)	969.00 (25.10%)	1,369.00 (24.20%)
162	home_shotsOnTarget_cat	3	2,011.00 (15.90%)	590.00 (18.60%)	777.00 (20.20%)	644.00 (11.40%)
169	away_shotsOnTarget_cat	4	2,114.00 (16.70%)	575.00 (18.10%)	664.00 (17.20%)	875.00 (15.50%)
176	home_total_assists_cat	1	4,382.00 (34.60%)	1,249.00 (39.40%)	1,099.00 (28.50%)	2,034.00 (36.00%)
180	away_total_assists_cat	1	4,335.00 (34.20%)	1,290.00 (40.70%)	1,498.00 (38.90%)	1,547.00 (27.40%)
184	home_corners_cat	5	1,758.00 (13.90%)	418.00 (13.20%)	526.00 (13.60%)	814.00 (14.40%)
198	home_Goals_cat	2	3,099.00 (24.40%)	637.00 (20.10%)	352.00 (9.10%)	2,110.00 (37.30%)
203	away_Goals_cat	0	4,031.00 (31.80%)	915.00 (28.80%)	0.00 (0.00%)	3,116.00 (55.10%)
208	home_total_blocked_shots_cat	2	2,318.00 (18.30%)	582.00 (18.30%)	685.00 (17.80%)	1,051.00 (18.60%)
222	home_total_saved_shots_cat	2	2,515.00 (19.80%)	645.00 (20.30%)	817.00 (21.20%)	1,053.00 (18.60%)

	Variables	Categories	Population	Category_1	Category_0	Category_2
229	away_total_saved_shots_cat	3	2,473.00 (19.50%)	606.00 (19.10%)	774.00 (20.10%)	1,093.00 (19.30%)
235	year	Mean (SD)	2,017.50 (2.10)	2,017.50 (2.10)	2,017.60 (2.10)	2,017.50 (2.00)
245	time_category	After Noon	5,278.00 (41.60%)	1,307.00 (41.20%)	1,659.00 (43.00%)	2,312.00 (40.90%)
321	home_discipline_score_avg	Mean (SD)	-19.20 (3.80)	-19.20 (3.70)	-19.10 (3.90)	-19.30 (3.70)
324	away_discipline_score_avg	Mean (SD)	-21.00 (4.00)	-20.90 (4.00)	-20.90 (4.10)	-21.10 (3.90)

62

```
array(['gameID', 'season', 'homeTeamID', 'awayTeamID', 'home_Goals',
       'away_Goals', 'home_GoalsHalfTime', 'away_GoalsHalfTime',
       'home_xGoals', 'home_shots', 'home_shotsOnTarget', 'home_deep',
       'home_ppda', 'home_fouls', 'home_corners', 'home_yellowCards',
       'home_redCards', 'home_total_assists', 'home_total_xAssists',
       'home_total_key_passes', 'home_total_xGoalsChain',
       'home_total_xGoalsBuildup', 'home_total_yellow_cards',
       'home_total_red_cards', 'home_total_blocked_shots',
       'home_total_saved_shots', 'away_xGoals', 'away_shots',
       'away_shotsOnTarget', 'away_deep', 'away_ppda', 'away_fouls',
       'away_corners', 'away_yellowCards', 'away_redCards',
       'away_total_assists', 'away_total_xAssists',
       'away_total_key_passes', 'away_total_xGoalsChain',
       'away_total_xGoalsBuildup', 'away_total_yellow_cards',
       'away_total_red_cards', 'away_total_blocked_shots',
       'away_total_saved_shots', 'home_redCards_binary',
       'away_redCards_binary', 'home_yellowCards_cat',
       'away_yellowCards_cat', 'home_shotsOnTarget_cat',
       'away_shotsOnTarget_cat', 'home_total_assists_cat',
       'away_total_assists_cat', 'home_corners_cat', 'home_Goals_cat',
       'away_Goals_cat', 'home_total_blocked_shots_cat',
       'home_total_saved_shots_cat', 'away_total_saved_shots_cat', 'year',
       'time_category', 'home_discipline_score_avg',
       'away_discipline_score_avg'], dtype=object)
```

We will add these variables to our variable selection table

	Variable	Univarable	Lasso
0	gameID	1	0
1	leagueID	0	0
2	season	1	0
3	date	0	0
4	homeTeamID	1	0
...	...	...	...
101	rolling_shot_rate_vs_blocks	0	0
102	discipline_matchup_score	0	0
103	interaction_away_goal_rate_vs_home_def	0	0
104	interaction_away_form_vs_home_cards	0	0
105	form_vs_discipline_gap	0	0

106 rows × 3 columns

Multivariable Analysis

	gameID	leagueID	season	date	homeTeamID	awayTeamID	home_Goals	away_Goals
5437	6185	5	2014	2014-08-08 19:30:00	177	161	2	2
5438	6186	5	2014	2014-08-09 20:00:00	162	164	3	3
5439	6187	5	2014	2014-08-09 20:00:00	209	165	0	3
5440	6188	5	2014	2014-08-09 20:00:00	169	175	0	2
5441	6189	5	2014	2014-08-09 20:00:00	160	180	0	0
...	...	...	...	...	...	...	...	...
12675	16131	5	2020	2021-05-23 19:00:00	168	166	1	2
12676	16132	5	2020	2021-05-23 19:00:00	177	176	1	2
12677	16133	5	2020	2021-05-23 19:00:00	163	235	2	0
12678	16134	5	2020	2021-05-23 19:00:00	175	181	0	1
12679	16135	5	2020	2021-05-23 19:00:00	225	179	1	1

12680 rows × 106 columns
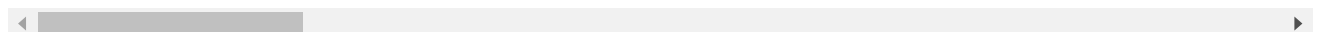
```
[(12680, 106), (12680, 1)]
```

Variable Selection using LASSO (L1 penalization)

```
array([False, False, False, False, False, False, False, False, False,
        True,  True, False, False,  True, False,  True, False,  True,
       False, False, False, False,  True, False, False, False, False,
       False, False, False, False,  True, False, False, False, False,
       False, False, False, False, False, False,  True, False, False,
       False, False, False, False,  True,  True, False,  True, False,
        True,  True,  True, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False])
```

Features in varSel but not in model:
{'date'}
Features in model but not in varSel:
set()

> ▼   Lasso

Lasso(alpha=0.01)

	Variable	Univarable	Lasso
10	home_xGoals	1	1
11	home_shots	1	1
14	home_ppda	1	1
16	home_corners	1	1
18	home_redCards	1	1
23	home_total_xGoalsBuildup	1	1
32	away_ppda	1	1
43	away_total_red_cards	1	1
50	home_shotsOnTarget_cat	1	1
51	away_shotsOnTarget_cat	1	1
53	away_total_assists_cat	1	1
55	away_corners_cat	0	1
56	home_Goals_cat	1	1
57	away_Goals_cat	1	1

14

	Variable	Univarable	Lasso
10	home_xGoals	1	1
11	home_shots	1	1
14	home_ppda	1	1
16	home_corners	1	1
18	home_redCards	1	1
23	home_total_xGoalsBuildup	1	1
32	away_ppda	1	1
43	away_total_red_cards	1	1
50	home_shotsOnTarget_cat	1	1
51	away_shotsOnTarget_cat	1	1
53	away_total_assists_cat	1	1
56	home_Goals_cat	1	1
57	away_Goals_cat	1	1

```
home_Goals_cat              0.503294
away_Goals_cat             -0.489818
away_ppda                  -0.021277
home_shotsOnTarget_cat      0.019747
home_shots                 -0.009302
home_total_xGoalsBuildup   -0.008499
home_xGoals                 0.008443
away_shotsOnTarget_cat     -0.008357
away_total_red_cards        0.008249
away_total_assists_cat     -0.007413
home_ppda                   0.003726
home_corners               -0.002685
home_redCards              -0.001923
away_corners_cat            0.000814
dtype: float64
```

Lasso Selected Feature Coefficients
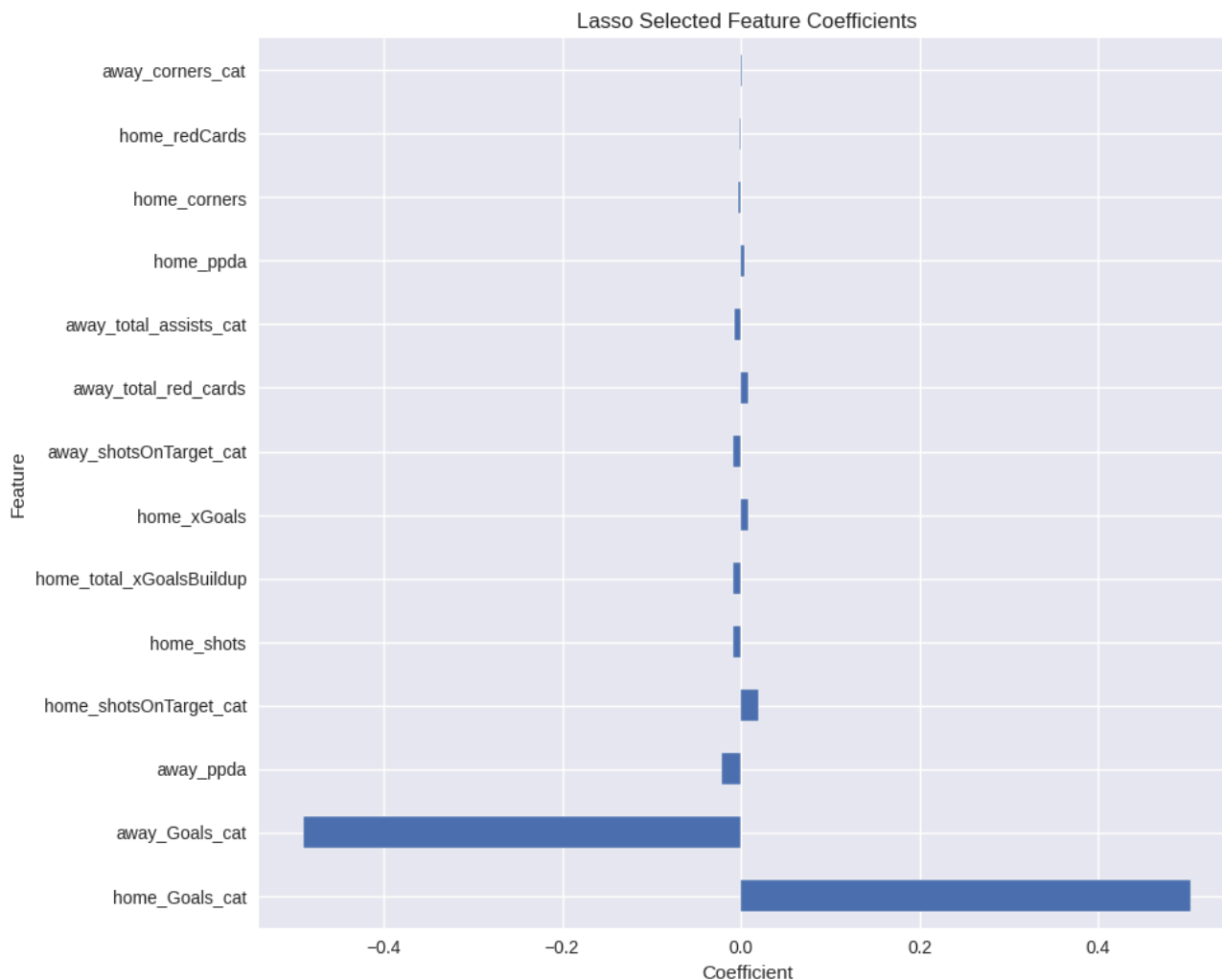
Variable Selection using Random Forest
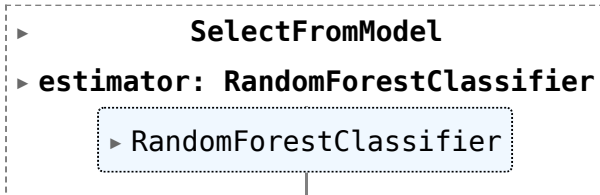
```
array([False, False, False, False, False,  True,  True,  True,  True,
        True, False, False, False, False, False, False, False, False,
        True,  True, False, False, False, False, False, False, False,
        True, False, False, False, False, False, False, False, False,
        True, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False,  True,  True, False,
       False,  True,  True, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False, False, False, False,
       False, False, False, False, False, False])
```

	Variable	Univarable	Lasso	RandomForest
6	home_Goals	1	0	1
7	away_Goals	1	0	1
8	home_GoalsHalfTime	1	0	1
9	away_GoalsHalfTime	1	0	1
10	home_xGoals	1	1	1
19	home_total_assists	1	0	1
20	home_total_xAssists	1	0	1
28	away_xGoals	1	0	1
37	away_total_assists	1	0	1
52	home_total_assists_cat	1	0	1
53	away_total_assists_cat	1	1	1
56	home_Goals_cat	1	1	1
57	away_Goals_cat	1	1	1

13

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
10	home_xGoals	1	1	1	0
53	away_total_assists_cat	1	1	1	0
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
10	home_xGoals	1	1	1	0
53	away_total_assists_cat	1	1	1	0
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

```
▸            SelectFromModel
▸ estimator: RandomForestClassifier
       ▸ RandomForestClassifier
```

```
home_Goals                    0.144157
away_Goals                    0.133237
home_Goals_cat                0.127784
away_Goals_cat                0.105205
home_total_assists            0.050917
away_total_assists            0.049632
home_total_assists_cat        0.042954
away_total_assists_cat        0.039263
away_GoalsHalfTime            0.023741
away_xGoals                   0.022401
home_GoalsHalfTime            0.020229
home_xGoals                   0.018936
home_total_xAssists           0.010480
away_shotsOnTarget            0.010151
home_shotsOnTarget            0.008964
home_total_xGoalsChain        0.008139
away_total_xGoalsBuildup      0.007158
away_total_xGoalsChain        0.007022
away_total_xAssists           0.006963
home_shotsOnTarget_cat        0.006163
dtype: float64
```


Top 20 Random Forest Feature Importances
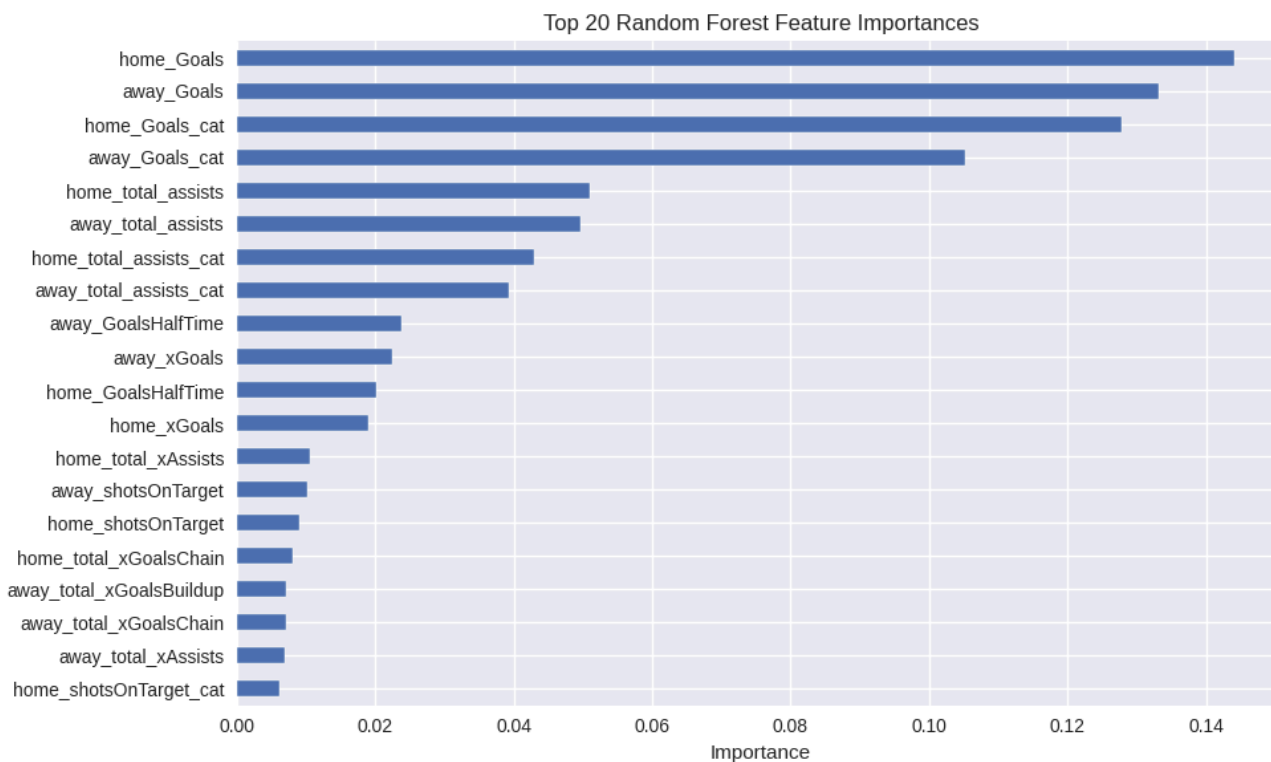
Variable Selection using Gradient Boosting classification

```
array([[False, False, False, False, False,  True,  True, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False,  True,  True, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False, False, False, False,
        False, False, False, False, False, False])
```

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
6	home_Goals	1	0	1	1
7	away_Goals	1	0	1	1
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

4

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
6	home_Goals	1	0	1	1
7	away_Goals	1	0	1	1
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

	Variable	Univarable	Lasso	RandomForest	GradienBoosting
56	home_Goals_cat	1	1	1	1
57	away_Goals_cat	1	1	1	1

```
▸            SelectFromModel
▸ estimator: GradientBoostingClassifier
    ▸ GradientBoostingClassifier
```

```
away_Goals                          0.270372
home_Goals                          0.254497
away_Goals_cat                      0.249122
home_Goals_cat                      0.225023
away_cumulative_goals_against       0.000377
awayTeamID                          0.000161
gameID                              0.000079
away_total_assists                  0.000040
away_total_assists_cat              0.000032
away_total_blocked_shots            0.000031
away_total_xGoalsBuildup            0.000026
away_avg_goals_for                  0.000025
month                               0.000022
form_vs_discipline_gap              0.000020
away_xGoals                         0.000019
away_GoalsHalfTime                  0.000017
home_ppda                           0.000015
rolling_xG_vs_shot_conceded         0.000014
home_GoalsHalfTime                  0.000012
away_total_xAssists                 0.000012
dtype: float64
```


Top 20 Gradient Boosting Feature Importances
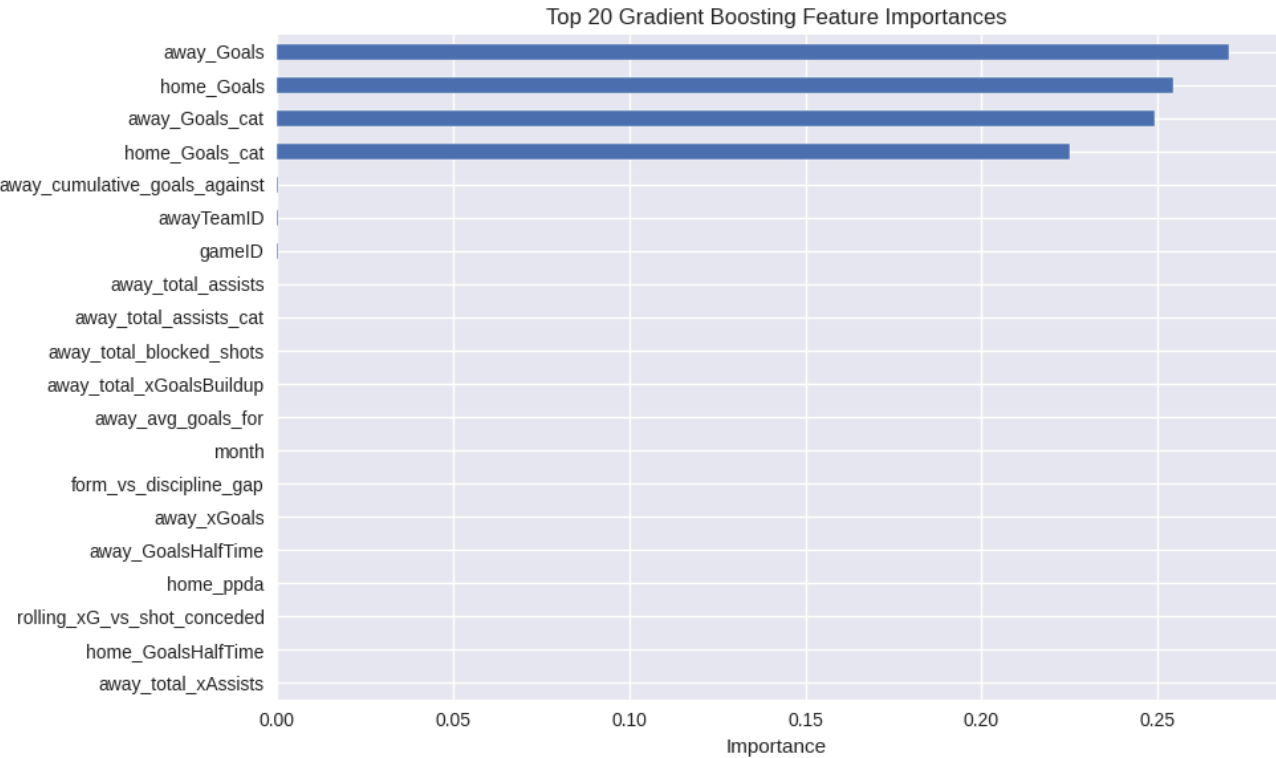
Variable Selection using SVM classification

```
array([ True, False, False,  True, False,  True,  True,  True, False,
       False,  True,  True, False,  True, False,  True, False, False,
        True, False, False,  True, False, False, False,  True, False,
        True, False,  True,  True,  True,  True,  True, False, False,
        True,  True,  True,  True, False,  True,  True,  True, False,
        True, False,  True, False,  True,  True,  True, False,  True,
        True,  True,  True,  True, False, False, False, False,  True,
       False,  True, False, False, False, False, False, False,  True,
       False, False, False,  True,  True, False, False,  True, False,
       False, False,  True, False,  True,  True, False,  True, False,
       False, False, False, False, False,  True,  True, False, False,
        True, False, False, False,  True,  True])
```

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SV
0	gameID	1	0	0	0	
4	homeTeamID	1	0	0	0	
6	home_Goals	1	0	1	1	
7	away_Goals	1	0	1	1	
8	home_GoalsHalfTime	1	0	1	0	
11	home_shots	1	1	0	0	
12	home_shotsOnTarget	1	0	0	0	
14	home_ppda	1	1	0	0	
16	home_corners	1	1	0	0	
19	home_total_assists	1	0	1	0	
22	home_total_xGoalsChain	1	0	0	0	
26	home_total_blocked_shots	1	0	0	0	
28	away_xGoals	1	0	1	0	
30	away_shotsOnTarget	1	0	0	0	
31	away_deep	1	0	0	0	
32	away_ppda	1	1	0	0	
33	away_fouls	1	0	0	0	
34	away_corners	1	0	0	0	
37	away_total_assists	1	0	1	0	
38	away_total_xAssists	1	0	0	0	
39	away_total_key_passes	1	0	0	0	
40	away_total_xGoalsChain	1	0	0	0	
42	away_total_yellow_cards	1	0	0	0	
43	away_total_red_cards	1	1	0	0	
44	away_total_blocked_shots	1	0	0	0	
46	home_redCards_binary	1	0	0	0	
48	home_yellowCards_cat	1	0	0	0	
50	home_shotsOnTarget_cat	1	1	0	0	
51	away_shotsOnTarget_cat	1	1	0	0	
52	home_total_assists_cat	1	0	1	0	
54	home_corners_cat	1	0	0	0	
55	away_corners_cat	0	1	0	0	
56	home_Goals_cat	1	1	1	1	
57	away_Goals_cat	1	1	1	1	
58	home_total_blocked_shots_cat	1	0	0	0	
63	month	0	0	0	0	

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SV
65	time_category	1	0	0	0	
72	away_cumulative_goal_diff	0	0	0	0	
76	away_avg_goals_against	0	0	0	0	
77	home_goal_per_shot	0	0	0	0	
80	away_goal_per_xg	0	0	0	0	
84	away_rolling_goal_per_xg	0	0	0	0	
86	away_blocked_shot_ratio	0	0	0	0	
87	home_shot_conceded_ratio	0	0	0	0	
89	home_discipline_score	0	0	0	0	
96	recent_form_diff	0	0	0	0	
97	interaction_goal_rate_vs_blocks	0	0	0	0	
100	rolling_xG_vs_shot_conceded	0	0	0	0	
104	interaction_away_form_vs_home_cards	0	0	0	0	
105	form_vs_discipline_gap	0	0	0	0	

50

```
home_Goals_cat           2.152787
away_Goals_cat           1.955288
away_Goals               1.454659
home_Goals               1.268327
away_shotsOnTarget_cat   0.038095
away_total_xAssists      0.034596
away_shotsOnTarget       0.033229
home_total_assists       0.027266
home_yellowCards_cat     0.025312
home_shotsOnTarget       0.024450
home_GoalsHalfTime       0.022891
away_fouls               0.022028
home_corners             0.019505
away_total_xGoalsChain   0.018256
home_shots               0.016529
away_total_assists       0.016480
home_total_xGoalsChain   0.015486
away_total_blocked_shots 0.014447
home_shotsOnTarget_cat   0.012255
away_rolling_goal_per_xg 0.012021
dtype: float64
```

Top 20 Feature Importances from SVM (L1)
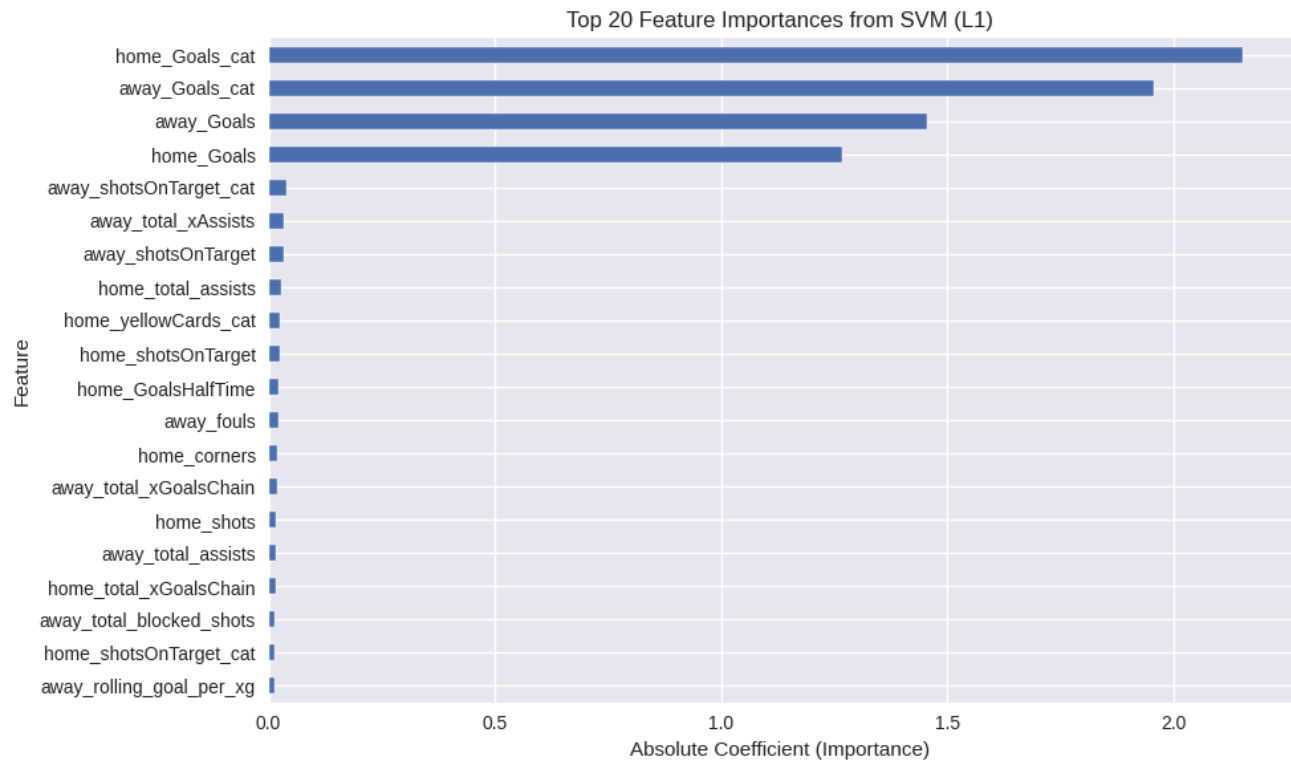
Summarization and Selection of Variables

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	
0	gameID	1	0	0	0	
1	leagueID	0	0	0	0	
2	season	1	0	0	0	
3	date	0	0	0	0	
4	homeTeamID	1	0	0	0	
...	...	...	...	...	...	
101	rolling_shot_rate_vs_blocks	0	0	0	0	
102	discipline_matchup_score	0	0	0	0	
103	interaction_away_goal_rate_vs_home_def	0	0	0	0	
104	interaction_away_form_vs_home_cards	0	0	0	0	
105	form_vs_discipline_gap	0	0	0	0	

106 rows × 7 columns

```
Sum
0    29
1    35
2    24
3    14
4     2
5     2
Name: Variable, dtype: int64
```

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SVM	Sum
56	home_Goals_cat	1	1	1	1	1	5
57	away_Goals_cat	1	1	1	1	1	5

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SVM	Sum
8	home_GoalsHalfTime	1	0	1	0	1	3
10	home_xGoals	1	1	1	0	0	3
11	home_shots	1	1	0	0	1	3
14	home_ppda	1	1	0	0	1	3
16	home_corners	1	1	0	0	1	3
19	home_total_assists	1	0	1	0	1	3
28	away_xGoals	1	0	1	0	1	3
32	away_ppda	1	1	0	0	1	3
37	away_total_assists	1	0	1	0	1	3
43	away_total_red_cards	1	1	0	0	1	3
50	home_shotsOnTarget_cat	1	1	0	0	1	3
51	away_shotsOnTarget_cat	1	1	0	0	1	3
52	home_total_assists_cat	1	0	1	0	1	3
53	away_total_assists_cat	1	1	1	0	0	3

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SVM	Sum
6	home_Goals	1	0	1	1	1	4
7	away_Goals	1	0	1	1	1	4
8	home_GoalsHalfTime	1	0	1	0	1	3
10	home_xGoals	1	1	1	0	0	3
11	home_shots	1	1	0	0	1	3
14	home_ppda	1	1	0	0	1	3
16	home_corners	1	1	0	0	1	3
19	home_total_assists	1	0	1	0	1	3
28	away_xGoals	1	0	1	0	1	3
32	away_ppda	1	1	0	0	1	3
37	away_total_assists	1	0	1	0	1	3
43	away_total_red_cards	1	1	0	0	1	3
50	home_shotsOnTarget_cat	1	1	0	0	1	3
51	away_shotsOnTarget_cat	1	1	0	0	1	3
52	home_total_assists_cat	1	0	1	0	1	3
53	away_total_assists_cat	1	1	1	0	0	3
56	home_Goals_cat	1	1	1	1	1	5
57	away_Goals_cat	1	1	1	1	1	5

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SVM	Sum
0	gameID	1	0	0	0	1	2
4	homeTeamID	1	0	0	0	1	2
6	home_Goals	1	0	1	1	1	4
7	away_Goals	1	0	1	1	1	4
8	home_GoalsHalfTime	1	0	1	0	1	3
9	away_GoalsHalfTime	1	0	1	0	0	2
10	home_xGoals	1	1	1	0	0	3
11	home_shots	1	1	0	0	1	3
12	home_shotsOnTarget	1	0	0	0	1	2
14	home_ppda	1	1	0	0	1	3
16	home_corners	1	1	0	0	1	3
18	home_redCards	1	1	0	0	0	2
19	home_total_assists	1	0	1	0	1	3
20	home_total_xAssists	1	0	1	0	0	2
22	home_total_xGoalsChain	1	0	0	0	1	2
23	home_total_xGoalsBuildup	1	1	0	0	0	2
26	home_total_blocked_shots	1	0	0	0	1	2
28	away_xGoals	1	0	1	0	1	3
30	away_shotsOnTarget	1	0	0	0	1	2
31	away_deep	1	0	0	0	1	2
32	away_ppda	1	1	0	0	1	3
33	away_fouls	1	0	0	0	1	2
34	away_corners	1	0	0	0	1	2
37	away_total_assists	1	0	1	0	1	3
38	away_total_xAssists	1	0	0	0	1	2
39	away_total_key_passes	1	0	0	0	1	2
40	away_total_xGoalsChain	1	0	0	0	1	2
42	away_total_yellow_cards	1	0	0	0	1	2
43	away_total_red_cards	1	1	0	0	1	3
44	away_total_blocked_shots	1	0	0	0	1	2
46	home_redCards_binary	1	0	0	0	1	2
48	home_yellowCards_cat	1	0	0	0	1	2
50	home_shotsOnTarget_cat	1	1	0	0	1	3
51	away_shotsOnTarget_cat	1	1	0	0	1	3
52	home_total_assists_cat	1	0	1	0	1	3
53	away_total_assists_cat	1	1	1	0	0	3

	Variable	Univarable	Lasso	RandomForest	GradienBoosting	SVM	Sum
54	home_corners_cat	1	0	0	0	1	2
55	away_corners_cat	0	1	0	0	1	2
56	home_Goals_cat	1	1	1	1	1	5
57	away_Goals_cat	1	1	1	1	1	5
58	home_total_blocked_shots_cat	1	0	0	0	1	2
65	time_category	1	0	0	0	1	2

```
pandas.core.frame.DataFrame
```

```
(12680, 19)
```

Setting proportion for Train, Test, Dev

```
Begining analysis...
Factorizing... please wait
********************
[]
********************
********************
['Unable to calcualte the Fisher exact test for variables away_total_red_cards and
split... The p-value may be incorrect']
------ Finished in 7.765838146209717econds -----

You got a perfectly balanced training and test datasets
```

```
(7608, 20)
```

Encoding Categorial Data

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 7608 entries, 8859 to 11119
Data columns (total 20 columns):
 #   Column                  Non-Null Count  Dtype
---  ------                  --------------  -----
 0   home_Goals              7608 non-null   int64
 1   away_Goals              7608 non-null   int64
 2   home_GoalsHalfTime      7608 non-null   int64
 3   home_xGoals             7608 non-null   float64
 4   home_shots              7608 non-null   float64
 5   home_ppda               7608 non-null   float64
 6   home_corners            7608 non-null   float64
 7   home_total_assists      7608 non-null   int64
 8   away_xGoals             7608 non-null   float64
 9   away_ppda               7608 non-null   float64
 10  away_total_assists      7608 non-null   int64
 11  away_total_red_cards    7608 non-null   int64
 12  home_shotsOnTarget_cat  7608 non-null   float64
 13  away_shotsOnTarget_cat  7608 non-null   float64
 14  home_total_assists_cat  7608 non-null   float64
 15  away_total_assists_cat  7608 non-null   float64
 16  home_Goals_cat          7608 non-null   float64
 17  away_Goals_cat          7608 non-null   float64
 18  gameresult              7608 non-null   int64
 19  split                   7608 non-null   float64
dtypes: float64(13), int64(7)
memory usage: 1.2 MB



Begining analysis...
Factorizing... please wait
*******************
[]
********************
********************
[]
------ Finished in 7.452263355255127econds -----

You got a perfectly balanced training and test datasets



(2536, 20)
```
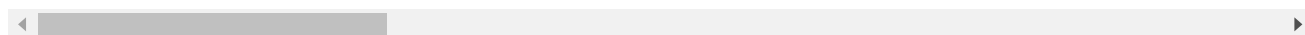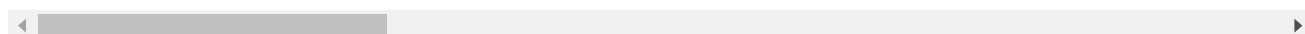
	home_Goals	away_Goals	home_GoalsHalfTime	home_xGoals	home_shots	home_ppda	h
9123	1	0	0	1.354740	10.000000	6.1111	
809	0	0	0	1.030380	19.000000	2.2500	
780	0	1	0	0.744526	15.000000	7.5758	
265	3	2	3	2.074880	11.000000	10.8571	
12190	1	1	0	1.754310	13.484856	3.3704	
...	...	...	...	...	...	...	
4626	1	4	1	1.495350	16.000000	10.6190	
11262	1	2	1	1.474980	13.000000	7.6000	
3731	1	3	1	0.600074	8.000000	10.0000	
439	2	0	1	0.569734	13.000000	10.9583	
1827	0	3	0	0.135570	3.000000	20.4286	

2536 rows × 20 columns

	home_Goals	away_Goals	home_GoalsHalfTime	home_xGoals	home_shots	home_ppda	h
10682	0	7	0	0.537278	11.0	8.2727	
10794	0	0	0	1.027300	14.0	18.6000	
2455	1	0	1	1.082450	9.0	9.6667	
6463	0	1	0	0.106979	2.0	10.0000	
168	0	0	0	0.973378	22.0	5.4706	
...	...	...	...	...	...	...	
3475	1	1	1	1.256860	15.0	4.4286	
1487	0	0	0	0.138495	4.0	6.5957	
4070	1	2	0	1.424000	16.0	4.9200	
4404	3	1	1	2.011290	7.0	11.8261	
9411	4	0	2	2.028830	17.0	12.7143	

2536 rows × 20 columns

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2536 entries, 10682 to 9411
Data columns (total 20 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   home_Goals             2536 non-null   int64
 1   away_Goals             2536 non-null   int64
 2   home_GoalsHalfTime     2536 non-null   int64
 3   home_xGoals            2536 non-null   float64
 4   home_shots             2536 non-null   float64
 5   home_ppda              2536 non-null   float64
 6   home_corners           2536 non-null   float64
 7   home_total_assists     2536 non-null   int64
 8   away_xGoals            2536 non-null   float64
 9   away_ppda              2536 non-null   float64
 10  away_total_assists     2536 non-null   int64
 11  away_total_red_cards   2536 non-null   int64
 12  home_shotsOnTarget_cat 2536 non-null   float64
 13  away_shotsOnTarget_cat 2536 non-null   float64
 14  home_total_assists_cat 2536 non-null   float64
 15  away_total_assists_cat 2536 non-null   float64
 16  home_Goals_cat         2536 non-null   float64
 17  away_Goals_cat         2536 non-null   float64
 18  gameresult             2536 non-null   int64
 19  split                  2536 non-null   object
dtypes: float64(12), int64(7), object(1)
memory usage: 416.1+ KB
```